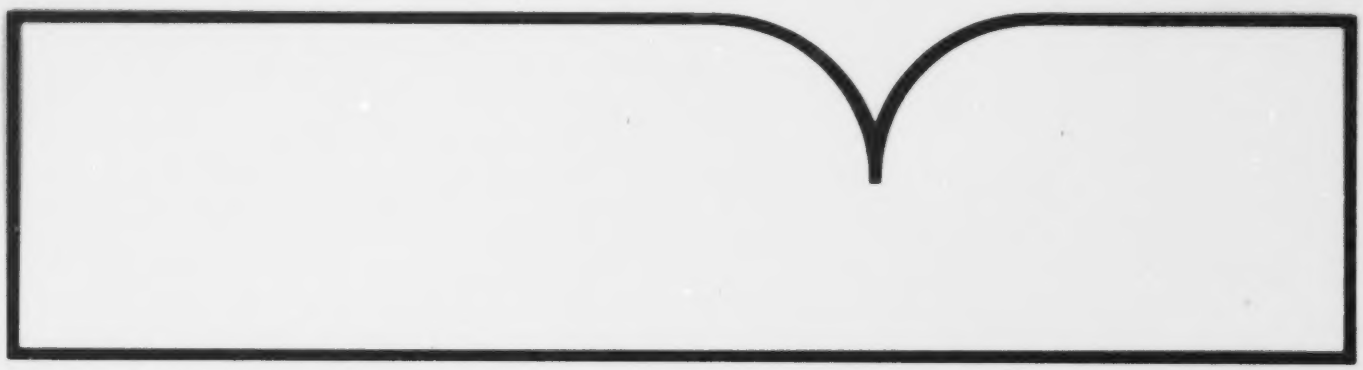
PB88-239157

Development and Evaluation of the Regional Oxidant Model for the Northeastern United States

(U.S.) Environmental Protection Agency
Research Triangle Park, NC

Aug 88



U.S. Department of Commerce
National Technical Information Service
NTIS

PB88-239157

EPA/600/D-88/169
August 1988

DEVELOPMENT AND EVALUATION OF THE REGIONAL OXIDANT MODEL FOR THE NORTHEASTERN UNITED STATES

by

K. L. Schere[1] and R. A. Wayland[2]

[1]National Oceanic and Atmospheric Administration, on assignment to U.S. Environmental Protection Agency, Research Triangle Park, NC 27711

[2]Computer Sciences Corporation, Research Triangle Park, NC 27709

EPA Project Officer
K. L. Schere

ATMOSPHERIC SCIENCES RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL
INFORMATION SERVICE
SPRINGFIELD, VA 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/D-88/169	2.	3. RECIPIENT'S ACCESSION NO. PB88-239157
4. TITLE AND SUBTITLE DEVELOPMENT AND EVALUATION OF THE REGIONAL OXIDANT MODEL FOR THE NORTHEASTERN UNITED STATES		5. REPORT DATE August 1988 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) K.L. Schere and R.A. Wayland		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS (SAME AS 12)		10. PROGRAM ELEMENT NO. FY-88 A101/B/63/21 4021-19 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Atmospheric Sciences Research Laboratory-RTP, NC Office of Research and Development US. Environmental Protection Agency Research Triangle Park, NC 27711		13. TYPE OF REPORT AND PERIOD COVERED IN-HOUSE 14. SPONSORING AGENCY CODE EPA/600/09

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The need for a comprehensive simulation model of photochemical smog on regional (1000 km) scales became apparent with the realization that the O_3 and oxidant pollution problem, first studied in the context of single urban areas, often extended over large geographical regions of the U.S., encompassing many urban areas. In response to this need EPA began a development effort toward a Regional Oxidant Model (ROM) nearly ten years ago. Today we have a second generation operational Model (ROM2) that continues to be developed, tested, and refined. The model is primarily designed for use in evaluating the effectiveness of various emission control strategies on the regional scale. The purpose of ~~this~~ the paper is to provide a brief overview of the ROM2 and to describe the first phase of a model evaluation study using data from the Northeast U.S.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS (This Report) UNCLASSIFIED	21. NO. OF PAGES 14
	20. SECURITY CLASS UNCLASSIFIED	22. PRICE A03 12.95

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

DEVELOPMENT AND EVALUATION OF THE REGIONAL OXIDANT MODEL FOR THE NORTHEASTERN UNITED STATES

K.L. Schere[1] and R.A. Wayland[2]

[1]National Oceanic and Atmospheric Administration, on assignment to the U.S. Environmental Protection Agency, Research Triangle Park, NC 27711 (U.S.A.)

[2]Computer Sciences Corporation, Research Triangle Park, NC 27709 (U.S.A.)

ABSTRACT

The second generation U.S. EPA Regional Oxidant Model (ROM2) has been developed over the last 10 years and is now operational. The 3-D grid model has been applied to the Northeast U.S. for a 50-day period in 1980. Model evaluation results show the ROM2 is performing well with respect to predicting the frequency distributions and spatial pattern of observed O_3 concentrations.

INTRODUCTION

The need for a comprehensive simulation model of photochemical smog on regional (1000 km) scales became apparent with the realization that the O_3 and oxidant pollution problem, first studied in the context of single urban areas, often extended over large geographical regions of the U.S., encompassing many urban areas. In response to this need EPA began a development effort toward a Regional Oxidant Model (ROM) nearly ten years ago. Today we have a second generation operational model (ROM2) that continues to be developed, tested, and refined (refs. 1-3). The model is primarily designed for use in evaluating the effectiveness of various emission control strategies on the regional scale. The purpose of this paper is to provide a brief overview of the ROM2 and to describe the first phase of a model evaluation study using data from the Northeast U.S.

MODEL OVERVIEW

The ROM has been designed to simulate most of the important chemical and physical processes responsible for the production of photochemically produced O_3 on scales of 1000 km, or several days of transport time. These processes include horizontal transport, atmospheric chemistry (using the Carbon Bond IV chemical mechanism, ref. 4), nighttime wind shear and turbulence episodes associated with the nocturnal jet, cumulus cloud effects on vertical mass transport and photochemical reaction rates, mesoscale vertical motions induced by terrain and the large scale flow, terrain effects on advection, diffusion

and deposition, sub-grid scale chemistry processes, emissions of natural and anthropogenic precursors, and dry deposition. They are mathematically simulated in the 3-D grid model with vertical resolution of 3 1/2 vertical layers including the boundary layer and the capping inversion or cloud layer. Horizontal resolution is approximately 18.5 km; however, the exact grid cell size varies somewhat over the model domain because the coordinate system used is based on latitude-longitude. Detailed descriptions of the theoretical basis and design of the ROM modeling system are available elsewhere (refs. 1-3).

The top three ROM layers are prognostic (predictive) and are free to locally expand and contract in response to changes in the physical processes occurring therein. The bottom layer is a shallow diagnostic surface layer designed to approximate the sub-grid scale effects on chemical reaction rates from a spatially heterogeneous emissions distribution. Layers 1 and 2 model the depth of the well-mixed layer during the day. In local areas where strong winds exist or the surface heat flux is weak, substantial wind shear can exist in the first few hundred meters above ground. This phenomenon can be treated separately in layer 1. Layer 3 represents the synoptic scale subsidence inversion, the base of which is typically 1-2 km above ground. Clouds, where they exist, are also treated in layer 3. At night, there is some readjustment in layer depths as the upper layers become decoupled from the lower surface inversion layer.

MODEL APPLICATION

In the current study the model has been applied to the Northeastern United States for the period July 12-August 31, 1980. Several regional photochemical smog episodes occurred during this period with measured O_3 concentrations as high as 300 ppb in the Northeast U.S. Emissions, air quality, and meteorological data bases were prepared for use by the ROM modeling system, which includes nearly 30 separate data preprocessing programs. In addition to the regular network of meteorological and air quality monitoring stations already in place during the period of interest, special aircraft and surface measurements were taken as part of an EPA-sponsored field study, the Northeast Regional Oxidant Study (ref. 5). A comprehensive anthropogenic source emissions inventory for hydrocarbons, NO_x, and CO was developed with a spatial and temporal resolution on the same scale as the ROM (ref. 6). In addition, an inventory of biogenically-produced hydrocarbon species was also developed (ref. 7) and included as part of the total emissions inventory.

The model simulation was started on a day with relatively clean tropospheric conditions. Initial model concentrations were set to values representative of these conditions, including NO_x=2 ppb, NMHC=15 ppbC, and O_3=35 ppb, with no spatial variation considered. Model simulation proceeded continuously through

the 50 day period, and the assumed initial concentration field was effectively swept out of the model domain by the fourth simulation day. Inflow boundary concentrations were updated during the simulation every 12 hours. Several remotely-located O_3 monitors were used to determine the temporal variations of inflow concentrations.

MODEL EVALUATION

Deterministic vs. non-deterministic modes

The O_3 surface monitoring networks of the portions of the U.S. and Canada included in the ROM simulation domain used here included 214 sites distributed throughout the model domain. The vast majority of the monitoring sites are located within or near urban and metropolitan areas. As a result there is good monitoring coverage in selected portions of the domain and poor spatial coverage in much of the remainder of the domain. Also, some of the urban-oriented monitors often show the effects of local NO_x scavenging of O_3. Nevertheless the large number of monitors and the extended simulation period provide useful data for performing an evaluation of the ability of the ROM to simulate O_3 concentrations over regional scales. We limit our discussion here to the performance of the model for predicting O_3 only near ground level.

Historically, model evaluation exercises have used the predictions of simulation models in a "deterministic" sense. That is, the model prediction for a given time and a given place is determined exactly by the data assimilated by the model and the model's component algorithms. This mode of operation does not explicitly consider model or data uncertainties. For the time and space scales used in urban-scale modeling, the deterministic mode may be appropriate. For the 1000 km regional scale, however, where multi-day transport is simulated, the deterministic mode becomes inappropriate. This is basically because the spatial scales of many key data elements are considerably coarser than the scales of the model. For instance, the upper-air meteorological network in North America that provides essential information for determining the transport component of the ROM has site locations separated by several hundred kilometers and soundings taken only every 12 hours. In the process of interpolating to the required spatial and temporal scales of the ROM, some of the determinism of the data set is lost to the various assumptions inherent in the interpolation method chosen; numerous interpolation methods exist each with its own set of assumptions. Lamb (refs. 8-9) maintains that more appropriate ways of interpreting the results of regional scale simulation models are in a "probabilistic" or "quasi-deterministic" sense.

In this study we use the "quasi-deterministic" mode, where model predictions and observations are not compared for a specific location and time, but rather

the comparisons are made for aggregate groups of receptor locations such that there exists a common characteristic uniting the members of each receptor group. For our analysis the groups were formed by analyzing the frequency distributions of daylight hourly O_3 concentrations during the summer of 1980 at the monitoring sites in our surface data base. Normalized frequencies of occurrence of O_3 concentrations between 5 and 20 ppb, 20 and 40 ppb, 40 and 80 ppb, and greater than 80 ppb were calculated for each monitoring site. A cluster analysis was performed using this data and 6 groups of receptor monitoring sites were selected based on their characteristic frequencies of observed O_3 concentrations. The groups include those with relatively high frequencies of >80 ppb O_3 (groups 1 and 2), those with more moderate levels (groups 3 and 4), one with unusually high frequencies of O_3 values <40 ppb (group 5), and one with relatively few low or high values (group 6).

Data preparation

Hourly averages of ROM O_3 predictions were interpolated from grid cell centers to the coordinates of each monitoring site location for the duration of the simulation period by a biquintic polynomial scheme. The observed O_3 monitoring data was passed through a 15-point high frequency filter designed to smooth data associated with those wavelengths shorter than the ROM could resolve based on its current grid resolution of 18.5 km. The net effect of this filter is to dampen very sharp peaks in a continuous time series of data. For O_3, a non-emitted species formed by atmospheric chemical reactions, sharp peaks in a temporal record are uncommon. Thus the filter's effect on the raw data is minimal. We expect to see a more pronounced effect of the filter on a primary species, such as NO_x, when we prepare that data set for comparison with ROM predictions later. Analyses for O_3 therefore were based on comparisons of the filtered observed and spatially interpolated predicted data for the aggregated monitoring sites within each of the groups discussed above. In addition to the group analyses, evaluation of the ROM's ability to replicate "plumes" of O_3 on the regional scale was tested by analyzing spatial patterns in the observed and predicted data. Here, the filtered observed and the hourly averaged grid predictions were used.

All ROM predicted data were taken from the first prognostic model layer above ground, typically extending from a base of 30-50 m above ground to a height of 100-300 m above ground during daytime conditions. The formulation of the shallow diagnostic surface layer in the model is not yet complete; therefore predictions for the surface, per se, were not available. By limiting our analyses to daytime conditions (0800 - 1900, Local Standard Time) when the lowest prognostic model layer well-represents the thermally mixed conditions near the surface, we minimize problems that sharp vertical gradients might cause

in the comparisons. Analyses for nighttime conditions or of sub-grid scale effects must await the final development of the ROM's layer 0 (surface layer).

Results

Table 1 presents a summary of the frequency distribution of O_3 concentrations for each group of receptors over the model simulation period. Both observed and predicted concentrations are summarized here. The number of monitoring sites within each group is also indicated.

TABLE 1

Frequency distribution of daylight observed and predicted O_3 concentrations within receptor groups for the period 14 July - 31 August 1980.

Group	Number of sites	Percent of daylight (08-19, LST) O_3 concentrations between: 5 - 20 ppb		21 - 40 ppb		41 - 80 ppb		> 80 ppb	
		Obs	Pred	Obs	Pred	Obs	Pred	Obs	Pred
1	35	8	1	19	7	39	66	34	26
2	39	15	1	23	9	41	69	21	21
3	64	16	1	28	9	46	79	10	11
4	54	22	1	37	12	36	80	5	7
5	20	42	2	38	16	19	73	1	9
6	2	0	0	26	4	63	93	11	3

These results seem to indicate that the ROM performs better at predicting O_3 concentrations greater than 40 ppb than it does at lower concentrations. Concentrations of 40 ppb and less were observed at the 6 groups of receptors from a high of 80% of the time (group 5) to a low of 26% of the time (group 6). The corresponding figures for predicted concentrations are 18% of the time (group 5) and 4% of the time (group 6). The model's performance for predicting concentrations >80 ppb appears quite good from the data in Table 1, with the possible exceptions of groups 5 and 6. The members of group 5 were chosen based on an unusually high frequency of O_3 concentrations less than 20 ppb. This indicates that these monitors were most likely influenced by near-source NO_x scavenging or other local processes that would cause ambient levels to drop considerably below background. Hence this is a pre-selected group of sites that are biased low in comparison to the model's expected predictions; the data in Table 1 appear to confirm the expectation.

Group 6 contains only 2 sites, those whose observed concentration distribution conforms to a pattern characteristic of remotely-located sites (i.e., relatively few values below tropospheric background levels). In fact, these sites are considerably distant from local source influences. One is located at Whiteface Mountain in northern New York and the other is at Long

Point Park in southern Ontario. The results for this group of monitors show the ROM to underpredict the occurrence of >80 ppb O_3. Small errors in the flow field can cause significant errors in the predicted trajectories of air parcels from particular source areas. These errors tend to worsen with increasing distance from the source. Since the remote sites are located at larger distances from source areas, one possible reason for the ROM's underprediction of higher concentrations at these sites is errors in the trajectories of air parcels affecting the sites from major source areas (see discussion of determinism in regional air quality models above).

The data in Table 1 also show that the ROM predicts the occurrence of moderate O_3 levels (40-80 ppb) considerably more frequently (66-93% of the time) than do the observations (19-63% of the time). This result follows from the model's underestimation of the frequency of low O_3 levels and is probably related to the local source influence at many of the urban-oriented monitoring sites.

Fig. 1 (a-f) presents plots of the observed cumulative frequency distribution of daytime O_3 concentrations versus the predicted distribution for each of the 6 receptor groups. The solid line in each figure represents the line of perfect agreement between the distributions and the dotted lines show the $\pm$10% error bound about the solid line. Each '+' symbol represents a percentile level in the frequency distribution, and the decile levels from 10% to 100% are marked by the integers 1,2,...,9,0. For example, in receptor group 1, approximately 80% of the observed O_3 concentrations and 90% of the predicted concentrations are less than 100 ppb. The plots confirm the model's tendency to overpredict at lower concentration values over all receptor groups. Group 1 contains the sites with the highest observed concentration values and the ROM shows a tendency of increasing underprediction with concentration level. The agreement between predictions and observations is best for groups 2-4 for moderate to high O_3 values, up through the 99.5% level of the frequency distributions. The maximum value of the distribution is consistently underpredicted by the model. Group 5, as discussed above, is an anomalous group. Fig. 1(e) confirms this anomaly by showing systematic overpredictions. Most groups show a more horizontal tendency in the distribution below about the 70-80% level, indicating that the range of the model predictions at lower concentrations is narrower than that of the observations.

Another key aspect of evaluating the ROM is its ability to replicate spatial patterns of observed pollutant concentrations. Fig. 2(a) shows the modeling domain used in this study. The maximum observed O_3 concentrations at the monitoring sites in the eastern portion of the domain for the episode that occurred during 20 July - 22 July are shown in Fig. 2(b), and the corresponding model predictions are contoured in Fig. 2(c). For this episode the model repli-

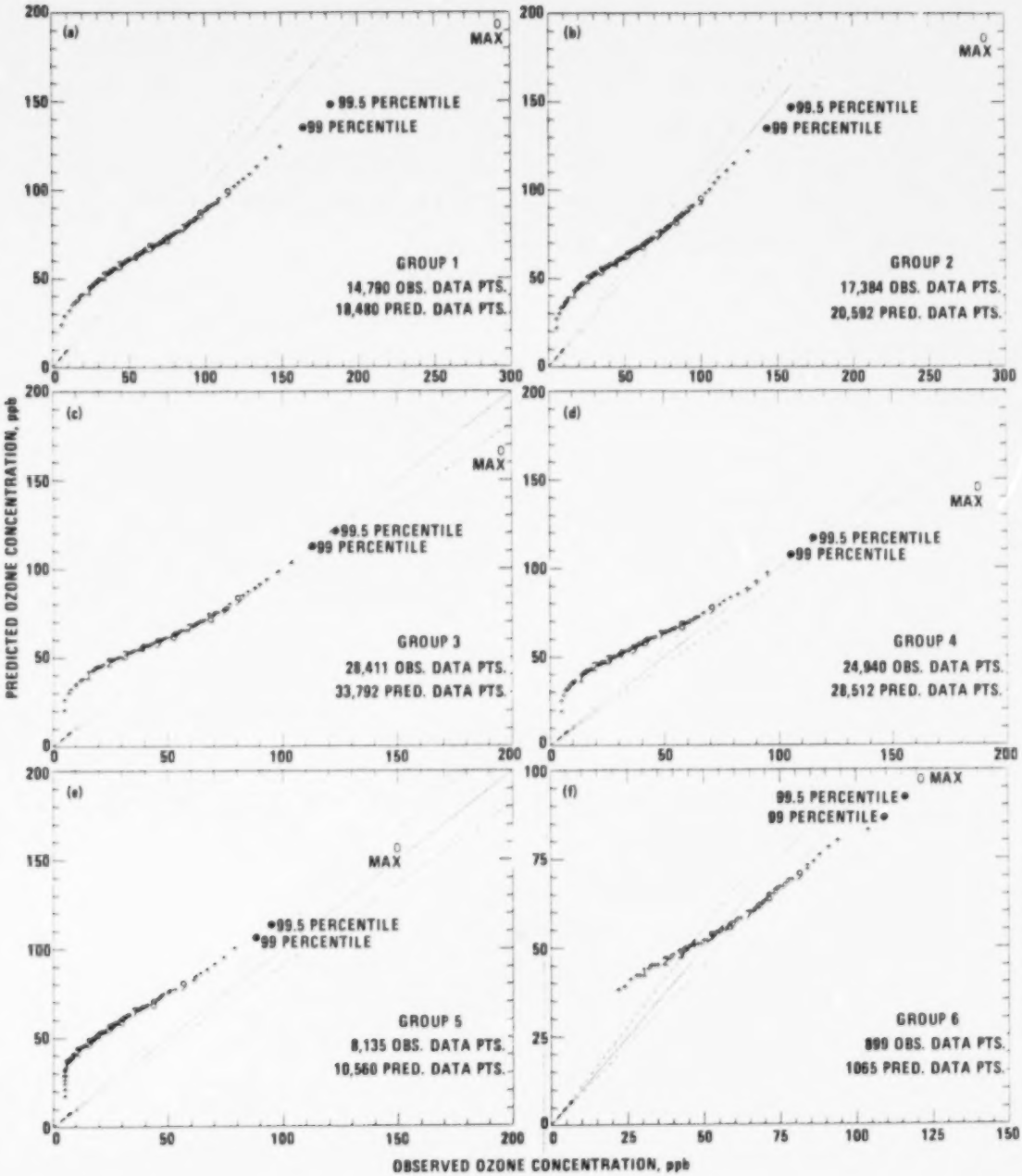


Fig. 1. Observed versus ROM-predicted cumulative frequency distributions of daytime (08-19 h, LST) hourly ozone concentrations at each of six groups of receptor locations over the period 14 July - 31 August 1980. Each percentile level in the distribution is shown and every tenth level is indicated by an integer.

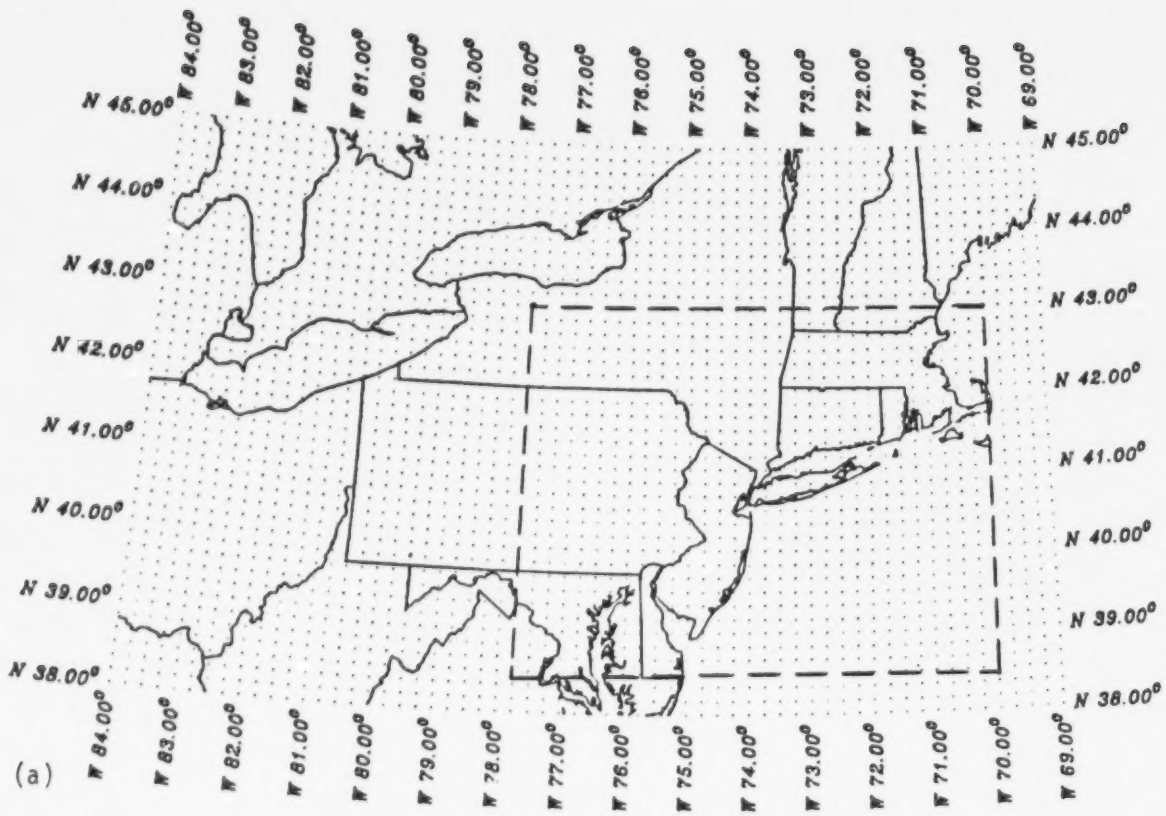


Fig. 2(a). Northeast U.S. ROM domain (dotted lines show area analyzed in Figs. 2(b) and 2(c)).

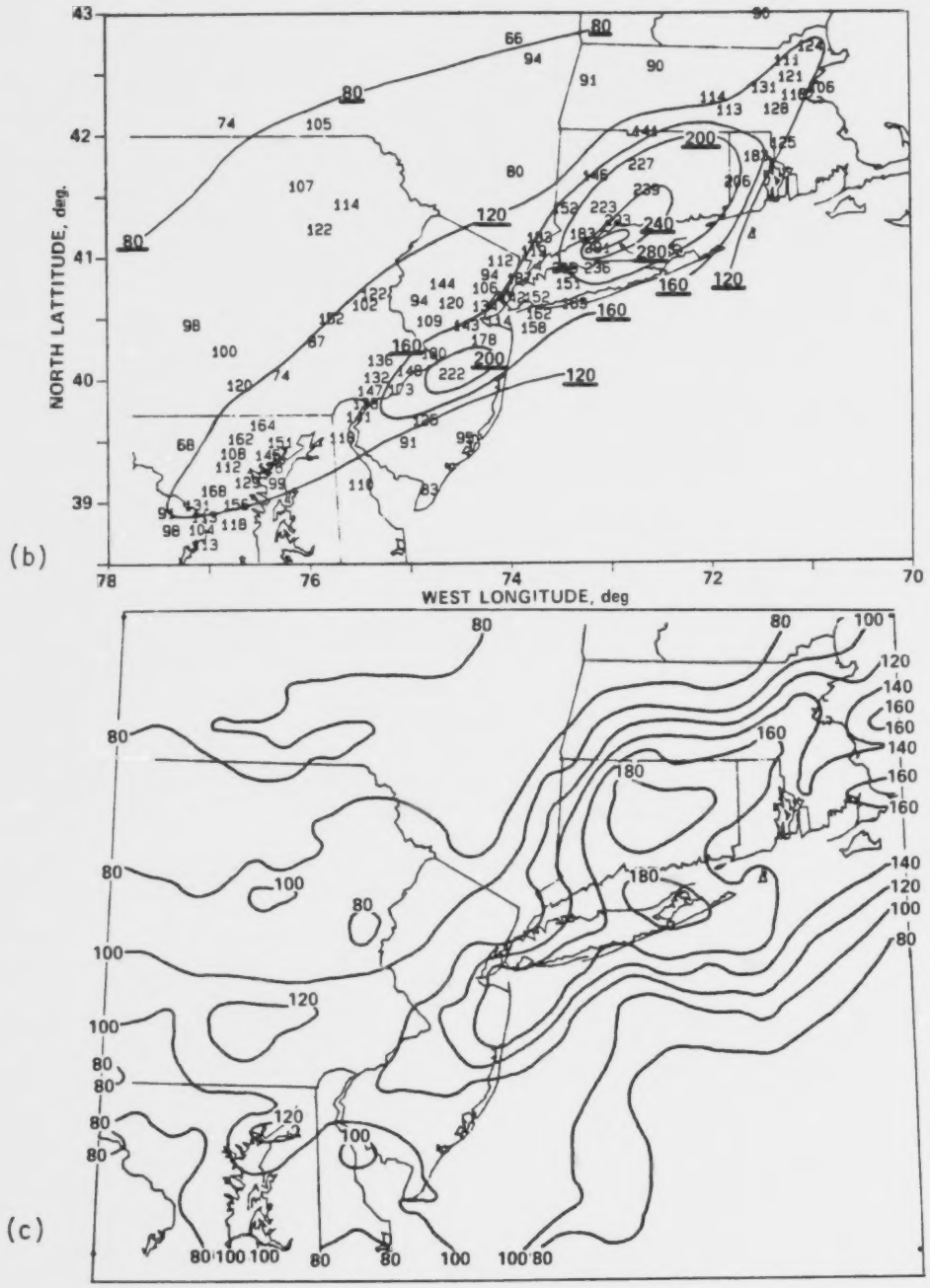


Fig. 2(b,c). Maximum observed hourly ozone concentrations (ppb) for the period 20-22 July 1980 (b), and contours of maximum predicted hourly ozone concentrations (ppb) for each ROM grid during the same period (c).

cated the spatial pattern of maximum O_3 quite well in the New England area. Portions of central Connecticut show >200 ppb O_3 both in the observations and the predictions. The ROM predicted another area of >200 ppb O_3 over eastern Long Island, although no monitoring sites existed there. In the corridor from Washington, D.C. to Philadelphia, the model underestimated maximum O_3 concentrations.

In summary, the ROM has been applied to the northeastern U.S. for a 50-day period during the summer of 1980. Preliminary model evaluation results show that the model reproduces the observed frequency distributions of O_3 concentrations best for moderate to high values. There appears to be a systematic tendency to overpredict the lowest observed values and to underpredict the highest values. The model replicated the spatial pattern of maximum O_3 concentrations in the New England area quite well for the 3-day episode containing some of the highest observed concentrations of the summer of 1980.

REFERENCES

1 R.G. Lamb, A Regional Scale (1000 km) Model of Photochemical Air Pollution. Part 1. Theoretical Formulation, EPA-600/3-83-035, U.S. EPA, Research Triangle Park, NC 27711, 1983, 239 pp.

2 R.G. Lamb, A Regional Scale (1000 km) Model of Photochemical Air Pollution. Part 2. Input Processor Network Design, EPA-600/3-84-085, U.S. EPA, Research Triangle Park, NC 27711, 1984, 298 pp.

3 R.G. Lamb and G.F. Laniak, A Regional Scale (1000 km) Model of Photochemical Air Pollution. Part 3. Tests of the Numerical Algorithms, EPA/600/3-85/037, U.S. EPA, Research Triangle Park, NC 27711, 1985, 265 pp.

4 G.Z. Whitten and M.W. Gery, Development of CBM-X Mechanisms for Urban and Regional AQSMs, EPA/600/3-86/012, U.S. EPA, Research Triangle Park, NC 27711, 1986, 160 pp.

5 W.M. Vaughan, Transport of Pollutants in Plumes and PEPES, EPA/600/3-85/033, U.S. EPA, Research Triangle Park, NC 27711, 1985, 60 pp.

6 D.A. Toothman, J.C. Yates and E.J. Sabo, Status Report on the Development of the NAPAP Emission Inventory for the 1980 Base Year and Summary of Preliminary Data, EPA-600/7-84-091, U.S. EPA, Research Triangle Park, NC 27711, 1984, 91 pp.

7 J.H. Novak and J.A. Reagan, A Comparison of Natural and Man-Made Hydrocarbon Emission Inventories Necessary for Regional Acid Deposition and Oxidant Modeling, 79th Annual Meeting of the APCA, Air Pollution Control Association, Pittsburgh, PA, 1986.

8 R.G. Lamb, Atmos. Environ., 18 (1984) 591-606.

9 R.G. Lamb and S.J. Hati, J. Cli. Appl. Met., 26 (1987) 837-846.